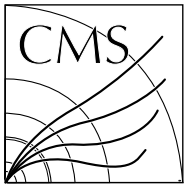
Available on CMS information server

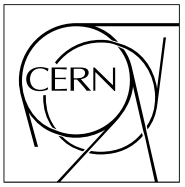
CMS NOTE 2006/109



The Compact Muon Solenoid Experiment

CMS Note

Mailing address: CMS CERN, CH-1211 GENEVA 23, Switzerland



8 June 2006

Charged MSSM Higgs Boson Observability in the $\mathrm{H}^{\pm} \to \mathrm{tb}$ Decay

S. Lowette[a)] and J. D'Hondt

Vrije Universiteit Brussel (VUB) – IIHE, Brussels, Belgium

P. Vanlaer

Université Libre de Bruxelles (ULB) – IIHE, Brussels, Belgium

Abstract

The discovery potential for a heavy charged MSSM Higgs boson is studied in the $\mathrm{H}^{\pm} \to \mathrm{tb}$ decay channel. Two strategies based on the tagging of either three or four b jets are considered. In both cases the main background from $\mathrm{t\bar{t}} + \mathrm{jets}$, with either additional b jets or jets mistagged as b, is found to be very large and kinematically similar to the signal. Systematic uncertainties on the background cross section and their impact on the search are also assessed. The discovery potential is investigated and no sensitivity for this channel is obtained in the MSSM parameter space. Assuming the same signal kinematics as in the MSSM, the cross section needed for a five–sigma discovery of a charged Higgs–boson signal is presented.

[a)] email: steven.lowette@cern.ch

1 Introduction

A natural way to extend the Higgs sector of the Standard Model (SM) is to add an extra complex Higgs doublet to the theory, giving rise to a general two Higgs doublet model (2HDM). In such a model five physical Higgs bosons remain after electroweak symmetry breaking [1]. Three of them are neutral: the scalars h and H and the pseudoscalar A; two other Higgs bosons are charged: the scalars $\mathrm{H}^{\pm}$. A particular example of a model containing such a 2HDM extension of the Higgs sector is the Minimal Supersymmetric Standard Model (MSSM) [2].

Two parameters describe the tree-level production and decay of the charged Higgs boson $\mathrm{H}^{\pm}$: the ratio of the vacuum expectation values of the two Higgs doublets $\tan\beta = v_2/v_1$ and the mass of the pseudoscalar Higgs boson m_{A}. The charged Higgs-boson mass is related to m_{A}, at tree level, by the relation

$$m_{\mathrm{H}^{\pm}} = \sqrt{m_{\mathrm{A}}^2 + m_{\mathrm{W}^{\pm}}^2}. \tag{1}$$

The branching fractions for the decay channels of the charged Higgs boson depend strongly on its mass. As shown in Figure 1, for $m_{\mathrm{H}^{\pm}} < m_{\mathrm{t}} + m_{\mathrm{b}}$, the $\mathrm{H}^{\pm} \to \tau\nu$ channel dominates. For larger masses, the channel

$$\mathrm{H}^{\pm} \to \mathrm{tb} \tag{2}$$

opens up. If the Higgs boson is produced through the inclusive channel $\mathrm{gb} \to \mathrm{tH}^{\pm}$, the decay (2) results in complex final states. From an experimental point of view, the most interesting final state is the semileptonic one,

$$\mathrm{gb} \to \mathrm{tH}^{\pm} \to \mathrm{ttb} \to \mathrm{W}^{+}\mathrm{W}^{-}\mathrm{bbb} \to \mathrm{qq}'\ell\nu\mathrm{bbb}, \tag{3}$$

since an isolated lepton (electron or muon) allows triggering and the branching fraction of this decay is large ($\sim 30\%$).

The potential of the decay channel (2) for large Higgs-boson masses at the LHC has been considered at parton level in several phenomenological studies [3–7]. These studies have shown the possibility of detecting the charged Higgs boson in certain regions of the $(\tan\beta, m_{\mathrm{A}})$ parameter space during the low-luminosity run of the LHC, with either three or four b-tagged jets. Crucial to these studies is a good b-identification capability to suppress the very large, and kinematically very similar $\mathrm{t\bar{t}}+\mathrm{jets}$ background. During this low luminosity period the LHC will operate at a luminosity $\mathcal{L} = 2 \times 10^{33}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$ and accumulate a total integrated luminosity of $L = \int \mathcal{L}\,\mathrm{d}t = 30\,\mathrm{fb}^{-1}$.

A fast-simulation study with parametrized detector performance requiring three b-tagged jets has been carried out previously by CMS [8]. This study took into account systematic uncertainties on the overall background cross section, and showed that no observability is left in the MSSM during the low luminosity phase of LHC.

In the analysis presented in this note, charged Higgs-boson detection is studied with full detector simulation in the

$$\mathrm{gg} \to \mathrm{tH}^{\pm}\mathrm{b} \to \mathrm{ttbb} \to \mathrm{W}^{+}\mathrm{W}^{-}\mathrm{bbbb} \to \mathrm{qq}'\ell\nu\mathrm{bbbb} \tag{4}$$

production mode, where a fourth b jet is resolved in the detector. Additionally, the production channel (3) using triple b tagging is reconsidered in the new fast-simulation framework of CMS. Both cases were only considered for final states with an isolated muon. Production of the $\mathrm{H}^{\pm}$ bosons through cascade decays of supersymmetric particles is not taken into account. It has been argued that for a heavy SUSY spectrum and $m_{\mathrm{H}^{\pm}} < 500\,\mathrm{GeV}/c^2$, the branching fraction for channel (2) is only slightly affected by SUSY decay channels opening up [9].

In Section 2 the signal and background simulations are described in detail. Next, in Section 3, the detector simulation and the event reconstruction and selection are discussed. Section 4 describes the analysis for the search with three b-tagged jets (3): the likelihood method to choose the optimal jet association, the charged Higgs-boson mass reconstruction, the background suppression, the determination of the significance and the discovery potential, and, finally, the impact of systematics on the background cross section prediction. In Section 5 the same analysis chain is summarized for the case (4) of four b-tagged jets.

2 Signal and Background Simulation

2.1 Signal simulation

At leading order (LO), events with three b-tagged jets originate form the production of the charged Higgs boson in the

$$\mathrm{gb} \to \mathrm{tH}^{\pm} \tag{5}$$

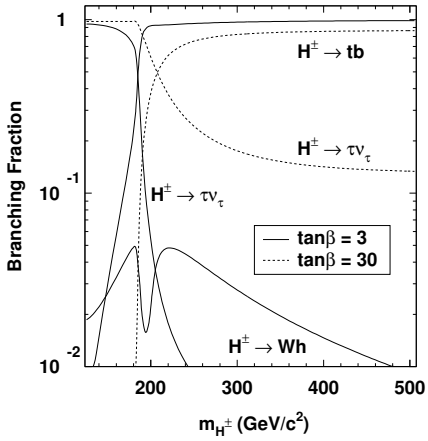


Figure 1: Charged Higgs boson branching fractions as a function of $m_{\mathrm{H}^\pm}$, obtained with HDECAY.

channel, with a top quark produced in association to the charged Higgs boson. The initial b quark is taken as a massless parton from the corresponding parton density in the proton. This description sums up large logarithms that arise from the phase-space region where the spectator b quark has small transverse momentum. At LO, events with four b-tagged jets originate from the

$$\mathrm{gg} \to \mathrm{tbH}^\pm \tag{6}$$

production process, which appropriately describes the kinematics of the spectator b quark, taken massive in the calculations. In this case the spectator b quark is produced via gluon splitting.

At the next-to-leading order (NLO), both processes (5) and (6) have to be properly combined. In the PYTHIA event generator [10], used to simulate the signal events, such a treatment at NLO is not possible. Recent developments allow to include a matching between both processes [11], but this was not considered in this study. The theoretical calculation of the signal cross section is performed at NLO, starting from the process (5) [12]. This calculation contains the process (6) as one of the NLO corrections. When calculating the cross section for both processes to all orders, however, one should obtain the same result, as they both describe the same physics. Therefore processes (5) and (6) were used to generate the final states (3) and (4) respectively, but the cross section was in each case rescaled to the NLO theoretical calculation for the $\mathrm{pp} \to \mathrm{tH}^\pm\mathrm{X}$ channel.

The signal cross section is sensitive to the two parameters $\tan\beta$ and $m_{\mathrm{H}^\pm}$. It is enhanced at small and large values of $\tan\beta$, with a minimum at $\tan\beta = \sqrt{m_\mathrm{t}/m_\mathrm{b}} \approx 6$. The cross section decreases rapidly with rising $m_{\mathrm{H}^\pm}$, losing one order of magnitude as $m_{\mathrm{H}^\pm}$ increases from $250\,\mathrm{GeV}/c^2$ to $500\,\mathrm{GeV}/c^2$, and when going from $\tan\beta = 30$ down to the minimum. The NLO $\mathrm{pp} \to \mathrm{tH}^\pm\mathrm{X}$ cross section as a function of $\tan\beta$ and $m_{\mathrm{H}^\pm}$ is shown in Figures 2 and 3.

The generation of the signal was performed with PYTHIA 6.223 for both processes (5) and (6), forcing the (2) decay. The branching fraction $\mathrm{BR}(\mathrm{H}^\pm \to \mathrm{tb})$ for this decay process, calculated with HDECAY 3.0 [13], ranges between $\sim 80\%$ for small $m_{\mathrm{H}^\pm}$ and large $\tan\beta$ and $\sim 100\%$ for large $m_{\mathrm{H}^\pm}$ and small $\tan\beta$, as shown in Figure 1. For both processes, 6 samples were generated at $\tan\beta = 30$ and masses $m_{\mathrm{H}^\pm}$ ranging from 263 to $506\,\mathrm{GeV}/c^2$ (corresponding to m_A between 250 and $500\,\mathrm{GeV}/c^2$). All possible decays of the two top quarks were considered.

2.2 Background simulation

The main background to charged Higgs boson production and decay through (3) and (4) is by far the Standard Model top-quark pair production with additional jets. Other potential multi-jet backgrounds are much smaller [14] and neglected in the following.

The LO background to process (3) comes from the SM process $\mathrm{pp} \to \mathrm{t\bar{t}b}$ and $\mathrm{pp} \to \mathrm{t\bar{t}} + \mathrm{jet}$, where in the latter the additional light quark or gluon jet is misidentified as a b jet. These background processes are not explicitly implemented in PYTHIA. With this generator only inclusive $\mathrm{t\bar{t}}$ production with additional jets from the parton shower can be performed, poorly describing the extra jet accompanying the $\mathrm{t\bar{t}}$ pair. To obtain a consistent description of the backgrounds at leading order, the event simulation should start from the hard interactions $\mathrm{pp} \to \mathrm{t\bar{t}b}$ and $\mathrm{pp} \to \mathrm{t\bar{t}j}$. This is implemented by using the matrix element generator MadGraph/MadEvent [15]. The

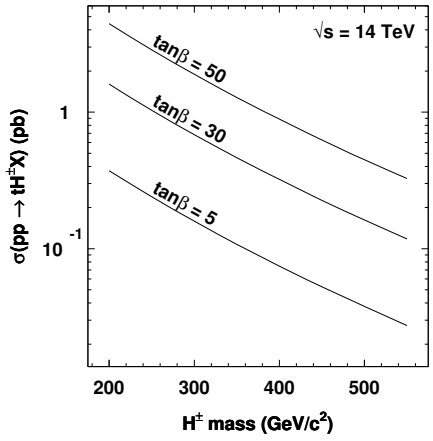


Figure 2: NLO cross section for the $\mathrm{pp} \to \mathrm{tH^{\pm}X}$ process as a function of $m_{\mathrm{H^{\pm}}}$.

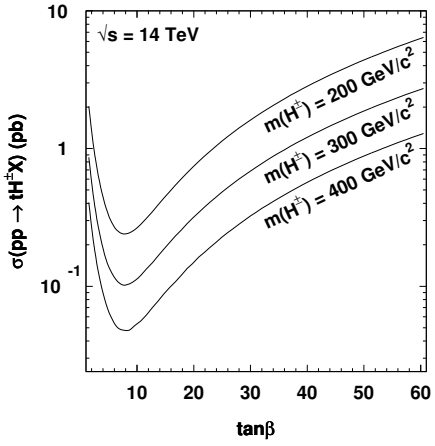


Figure 3: NLO cross section for the $\mathrm{pp} \to \mathrm{tH^{\pm}X}$ process as a function of $\tan\beta$.

events were generated requiring the extra parton to have transverse momentum $p_T > 10\,\mathrm{GeV}/c$ and pseudorapidity $|\eta| < 2.5$, in order to keep the cross section finite. These cuts resulted in a total cross section of $678\,\mathrm{pb}$. After the simulation of the hard interaction, the events were interfaced to PYTHIA for parton showering, decay and hadronization. In the following, this background is called the $\mathrm{t\bar{t}b/t\bar{t}j}$ background.

The LO background for process (4) consists of the irreducible $\mathrm{pp} \to \mathrm{t\bar{t}b\bar{b}}$ process and of the reducible $\mathrm{pp} \to \mathrm{t\bar{t}jj}$ process, where two non-b jets are misidentified as b jets. Both backgrounds have been produced using the CompHEP v.41.10 generator [16]. The generator level cuts $p_T > 15\,\mathrm{GeV}/c$ and $|\eta| < 3$ were applied on the partons produced in association with the $\mathrm{t\bar{t}}$ pair (b or non b). An additional separation cut $\Delta R > 0.3$ was imposed between these partons. This results in a cross section of $3.285\,\mathrm{pb}$ for the $\mathrm{pp} \to \mathrm{t\bar{t}b\bar{b}}$ process and $507.8\,\mathrm{pb}$ for $\mathrm{pp} \to \mathrm{t\bar{t}jj}$ production.

Double counting of events can occur between the $\mathrm{pp} \to \mathrm{t\bar{t}b\bar{b}}$ and $\mathrm{pp} \to \mathrm{t\bar{t}jj}$ processes when a gluon in the latter process splits into a $\mathrm{b\bar{b}}$ pair. To avoid this in a rigourous way, a NLO event generator should be used to simultaneously simulate both processes. Such a generator was not available at the time of writing, therefore double counting was avoided by rejecting $\mathrm{t\bar{t}jj}$ events with a $\mathrm{b\bar{b}}$ pair with $p_T > 15\,\mathrm{GeV}/c$, $|\eta| < 3$ and $\Delta R > 0.3$.

Another problem of double counting that occurs in multi-jet processes at leading order involves the transition between the matrix element for the hard interaction and the parton shower. A jet-matching technique between matrix element and parton-shower jets that allows to more reliably handle exclusive n-jet final states has been deployed using a combination of the ALPGEN v2.05 [17] and PYTHIA Monte-Carlo programs. Applying the same generator cuts on the CompHEP $\mathrm{t\bar{t}jj}$ sample as on the inclusive $\mathrm{t\bar{t}jj}$ ALPGEN sample, one obtains a cross section correction factor of 0.64 for the CompHEP cross section, which is applied throughout this note. The remaining kinematical differences between the two samples have been neglected.

No constraints were applied for the decay of background events. As a consequence, all $\mathrm{W^{\pm}}$ decay channels were considered.

3 Simulation, Reconstruction and Selection

3.1 Detector simulation and reconstruction of physics objects

A GEANT4 [18] based CMS detector simulation was used for the signal process (4) and its corresponding backgrounds. On average 3.5 minimum bias events were superimposed on each event, to account for the event pile-up expected during the low luminosity phase of the LHC. In the case of the signal process (3) and its corresponding background, the event generation was followed by the parametrized detector simulation. The reconstruction and analysis code was designed in such a way that it is identical for both the full and the parametrized detector simulation, except for the Level-1 trigger which is not implemented in the fast simulation.

When reconstructing isolated muons, muon candidates are first taken from an algorithm which combines track segments in the muon chambers with information from the tracker. To identify the muons that result from $\mathrm{t} \to \mathrm{W^{\pm}} \to \mu^{\pm}$ decays, a discriminator ($\mathcal{L}_{\mathrm{muon}}$) is calculated for each muon candidate [19]. It combines iso-

lation variables from calorimeter and tracker information, with the momentum of the muon and the association significance to the primary vertex into a global discriminator. Among all muon candidates in the event, the muon candidate with the largest discriminator $\mathcal{L}_{\mathrm{muon}}$, fulfilling the condition $\mathcal{L}_{\mathrm{muon}} > 0.01$, is identified as the one from the $\mathrm{W}^{\pm}$ decay.

Jets are reconstructed using an iterative cone clustering algorithm with a cone opening angle $\Delta R = 0.5$ and a cut on the E_T of the seeds at $2\,\mathrm{GeV}$. Calorimeter towers were used as input to the clustering algorithm. Variable tower thresholds account for larger energy deposits from the underlying event in the forward region [20]. The jet clustering excludes the muon. Jet energy scale corrections are applied to the reconstructed jets using Monte Carlo jet-energy corrections [20]. Finally, jets emerging from the pile-up collisions are vetoed using a track-based method. A jet is associated with the primary vertex if $\beta = \sum_{\mathrm{track,PV,jet}} p_T / \sum_{\mathrm{track,jet}} p_T > 0.04$, where the sum in the denominator is over the p_T of all tracks in the jet, while the sum in the nominator runs only over those tracks associated with the primary vertex.

The identification of b jets is performed using a likelihood-based algorithm that exploits secondary vertex reconstruction, distinguishing between several vertex categories [21]. The output of the algorithm is a combined b-jet discriminator, ζ_{b}, shown in Figure 4 for the different jet flavours. A b-jet probability P_{b} is derived from this discriminator. It is defined as the ratio $S/(S+B)$, where S is the number of b jets in a sample at a given ζ_{b}, and B the corresponding number of non-b jets. The distribution of P_{b} as a function of ζ_{b} is shown in Figure 5.

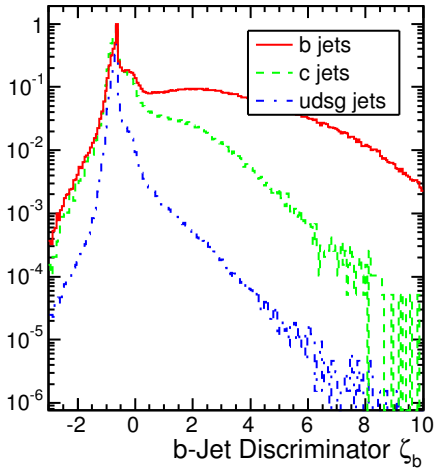


Figure 4: Distribution of the b-jet discriminator ζ_{b} for the different jet flavours.

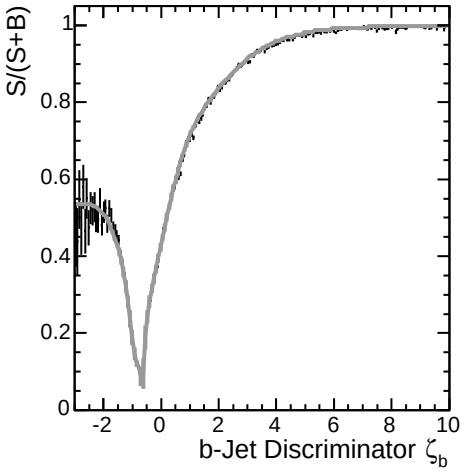


Figure 5: The b-jet probability P_{b} as a function of the b-jet discriminator ζ_{b}.

The final part of the event reconstruction deals with the missing transverse energy. It is calculated as the vectorial sum of the transverse components of all energy deposits in the calorimeters and corrected for the transverse energy carried by the muons.

3.2 Event selection and reconstruction

The current High Level Trigger (HLT) threshold for the single muon trigger ($p_T(\mu) > 19\,\mathrm{GeV}/c$) [22] is applied for both final states (3) and (4). Events passing this trigger criterion are required to have

- at least one muon with $p_T > 20\,\mathrm{GeV}/c$ and $|\eta| < 2.5$;
- at least five or six jets with $p_T > 25\,\mathrm{GeV}/c$ and $|\eta| < 2.5$, for processes (3) and (4), respectively;
- at least three or four b–tagged jets, for processes (3) and (4), respectively. The ζ_{b} thresholds to tag a jet as b are fixed at 1.0 and 0.5, respectively, the latter being softer due to the reduced statistics of final state (4).

The relative efficiencies of these cuts and the remaining number of events after the complete event selection are summarized in Table 1 and 2 for the final states (3) and (4) respectively, for $\tan\beta = 30$ and an integrated luminosity of $30\,\mathrm{fb}^{-1}$.

Table 1: Selection yield for $\tan\beta = 30$ and an integrated luminosity of $30\,\mathrm{fb}^{-1}$.

$\tan\beta = 30$ $30\,\mathrm{fb}^{-1}$	$\mathrm{t\bar{t}b/t\bar{t}j}$	$\mathrm{gb \to tH^{\pm}}$ $m_{\mathrm{H^{\pm}}} = 263\,\mathrm{GeV}/c^2$	$\mathrm{gb \to tH^{\pm}}$ $m_{\mathrm{H^{\pm}}} = 311\,\mathrm{GeV}/c^2$	$\mathrm{gb \to tH^{\pm}}$ $m_{\mathrm{H^{\pm}}} = 359\,\mathrm{GeV}/c^2$	$\mathrm{gb \to tH^{\pm}}$ $m_{\mathrm{H^{\pm}}} = 408\,\mathrm{GeV}/c^2$	$\mathrm{gb \to tH^{\pm}}$ $m_{\mathrm{H^{\pm}}} = 457\,\mathrm{GeV}/c^2$	$\mathrm{gb \to tH^{\pm}}$ $m_{\mathrm{H^{\pm}}} = 506\,\mathrm{GeV}/c^2$
cross section × BR (pb)	678	0.850	0.570	0.377	0.251	0.169	0.116
# events before cuts	20.3×10^6	25 489	17 088	11 319	7 529	5 063	3 472
single muon HLT	17%	16%	16%	16%	16%	16%	16%
1 muon	95%	95%	95%	95%	96%	96%	96%
5 jets	18%	35%	42%	44%	46%	49%	51%
3 b-tagged jets	6%	27%	29%	30%	32%	31%	29%
# remaining events	32 880	364	314	230	171	116	80

Table 2: Selection yield for $\tan\beta = 30$ and an integrated luminosity of $30\,\mathrm{fb}^{-1}$.

$\tan\beta = 30$ $30\,\mathrm{fb}^{-1}$	$\mathrm{t\bar{t}b\bar{b}}$	$\mathrm{t\bar{t}jj}$	$\mathrm{gg \to tbH^{\pm}}$ $m_{\mathrm{H^{\pm}}} = 263\,\mathrm{GeV}/c^2$	$\mathrm{gg \to tbH^{\pm}}$ $m_{\mathrm{H^{\pm}}} = 311\,\mathrm{GeV}/c^2$	$\mathrm{gg \to tbH^{\pm}}$ $m_{\mathrm{H^{\pm}}} = 359\,\mathrm{GeV}/c^2$	$\mathrm{gg \to tbH^{\pm}}$ $m_{\mathrm{H^{\pm}}} = 408\,\mathrm{GeV}/c^2$	$\mathrm{gg \to tbH^{\pm}}$ $m_{\mathrm{H^{\pm}}} = 457\,\mathrm{GeV}/c^2$	$\mathrm{gg \to tbH^{\pm}}$ $m_{\mathrm{H^{\pm}}} = 506\,\mathrm{GeV}/c^2$
cross section × BR (pb)	2.386	235.8	0.850	0.570	0.377	0.251	0.169	0.116
# events before cuts	71 580	7.07×10^6	25 489	17 088	11 319	7 529	5 063	3 472
single muon HLT	19%	19%	13%	13%	13%	13%	13%	13%
1 muon	96%	97%	96%	95%	97%	97%	97%	97%
6 jets	19%	23%	19%	23%	25%	26%	28%	31%
4 b-tagged jets	7%	0.55%	6%*	5%*	7%*	7%*	5%*	6%*
# remaining events	179	1 623	37	24	25	18	9	8

* large uncertainty due to low number of simulated and selected events

A kinematic fit which imposes both the $\mathrm{W}^{\pm}$-boson and both the top-quark mass constraints is applied. It is based on a linearized iterative least-square method with Lagrange multipliers [23]. The mass of the jets is not conserved in the fit, while the muon and the neutrino masses are fixed. In the fit, the objects' kinematics are parametrized by the inverse momentum $1/p$, the polar angle θ and the azimuthal angle ϕ. Jets are further described by the energy parameter $d = E_{\mathrm{fitted}}/E_{\mathrm{measured}}$. The diagonal elements of the covariance matrices, corresponding to the resolutions of the jets and lepton in the chosen parametrization, are determined as a function of the transverse momentum p_T in a $\mathrm{t\bar{t}} \to \mathrm{b\bar{b}qq'}\mu\nu$ sample. The neutrino is treated separately. Its transverse momentum is identified with the transverse missing momentum vector. Its resolution in θ and $1/p$ are chosen to be large enough to give full freedom in the fit to the unknown z component. The resolution of the angle ϕ of the neutrino object is taken as a constant, 0.54, as obtained from a study of a $\mathrm{t\bar{t}} \to \mathrm{b\bar{b}qq'}\mu\nu$ sample.

The kinematic fitting procedure was applied for all possible jet associations. As a result of the fit a χ^2 probability is returned if the fit converged, along with the parameter vectors of the fitted objects. Events for which the fit did not converge for any jet association were discarded.

4 Charged Higgs-Boson Search in the Three b-Jet Channel

4.1 Selection of the optimal jet association

Five jets are present in the final state (3), of which three are b jets. Out of the 60 possible jet associations to the final state quarks, the b-tagging information reduces the number of combinations to six. Several observables were identified to choose the best solution:

- observables associated with the output of the kinematic fit: the χ^2 probability of the fit (Figure 6) and the mass difference $m(\mathrm{t}_{\mathrm{hadr,fit}}) - m(\mathrm{t}_{\mathrm{hadr,rec}})$ of the hadronically decaying top quark after and before the fit (Figure 7);
- the transverse momentum and pseudorapidity of the b jet from the $\mathrm{H}^{\pm}$ decay: $p_T(\mathrm{b}_{\mathrm{H}^{\pm}})$ (Figure 8) and $|\eta(\mathrm{b}_{\mathrm{H}^{\pm}})|$ (Figure 9);
- the combined b-identification probability

$$P_{3\mathrm{b}} = \frac{P_{\mathrm{b}}(\mathrm{b}_1) + P_{\mathrm{b}}(\mathrm{b}_2) + P_{\mathrm{b}}(\mathrm{b}_3)}{P_{\mathrm{b}}(\mathrm{b}_1) + P_{\mathrm{b}}(\mathrm{b}_2) + P_{\mathrm{b}}(\mathrm{b}_3) + P_{\mathrm{b}}(\mathrm{j}_1) + P_{\mathrm{b}}(\mathrm{j}_2)},$$

  obtained from the b-jet probabilities of the three b jets and of the two jets from the $\mathrm{W}^{\pm}$ decay (Figure 10.

The jet associations used to obtain the distributions for these observables include the ones disfavoured by the b tagging, except for the associations for which the kinematic fit fails. For each of the observables, a corresponding likelihood ratio function $\mathcal{L}_i$ was obtained by calculating bin-by-bin the ratio $S/(S+B)$, where S and B denote respectively the number of correct and wrong jet associations, and by fitting the resulting distribution. The individual $\mathcal{L}_i$ distributions are shown in Figures 6 through 10.

For each jet association a combined likelihood ratio value is calculated as

$$\mathcal{L}_{\mathrm{sol}} = \prod_i \mathcal{L}_i. \tag{7}$$

The event solution with the highest value for $\mathcal{L}_{\mathrm{sol}}$ is chosen as the best possible jet association. This choice results in all jets correctly associated in 13% of the events.

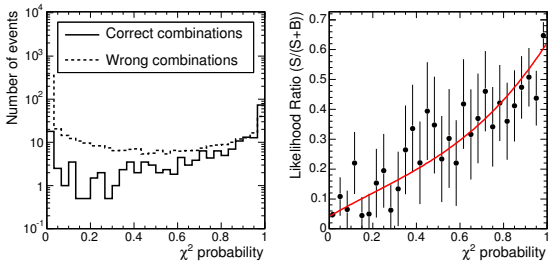


Figure 6: χ^2 probability for $m_{\mathrm{H}^{\pm}} = 311\,\mathrm{GeV}$.

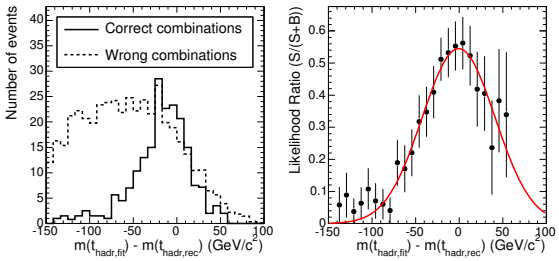


Figure 7: $m(\mathrm{t}_{\mathrm{hadr,fit}}) - m(\mathrm{t}_{\mathrm{hadr,rec}})$ for $m_{\mathrm{H}^{\pm}} = 311\,\mathrm{GeV}$.

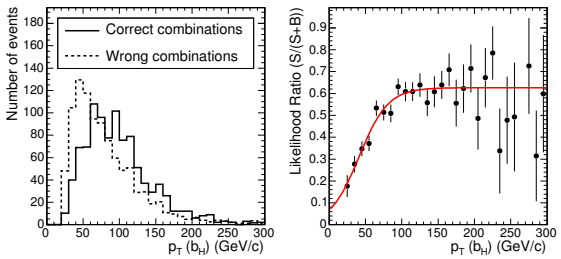


Figure 8: $p_T(\mathrm{b}_{\mathrm{H}^{\pm}})$ for $m_{\mathrm{H}^{\pm}} = 311\,\mathrm{GeV}$.

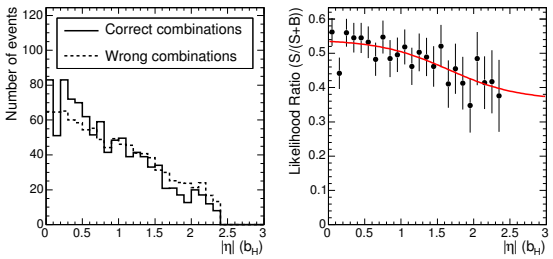


Figure 9: $|\eta(\mathrm{b}_{\mathrm{H}^{\pm}})|$ for $m_{\mathrm{H}^{\pm}} = 311\,\mathrm{GeV}$.

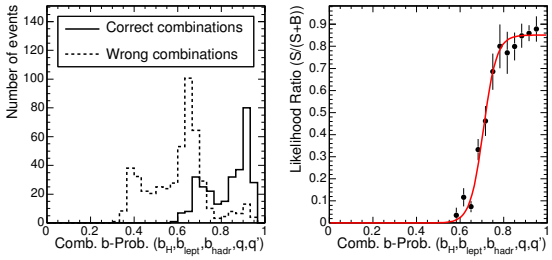


Figure 10: Combined b probability $P_{3\mathrm{b}}$ for $m_{\mathrm{H}^\pm} = 311\,\mathrm{GeV}$. (The peak structure is a consequence of the large b-probability separation between b and non-b quarks)

4.2 Mass reconstruction

As the longitudinal momentum of the neutrino is determined from the kinematic fit, it is possible to reconstruct the charged Higgs boson mass. It was shown [23] that the kinematic fit will in addition reduce the width of the reconstructed mass peak. An ambiguity in the reconstruction remains however, as it is not possible to know which top quark the additional b jet should be combined with.

Figures 11 and 12 present the reconstructed Higgs-boson mass for the correct jet associations and for the chosen jet associations, respectively. Hadronic decays of the top quarks are considered. Figures 13 and 14 present the same distributions for leptonic decays of the top quark.

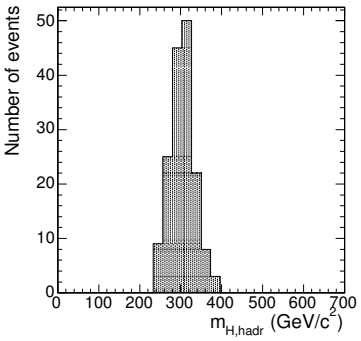


Figure 11: Reconstructed $\mathrm{H}^\pm$ mass with hadronically decaying top quark for correct jet associations ($m_{\mathrm{H}^\pm} = 311\,\mathrm{GeV}$).

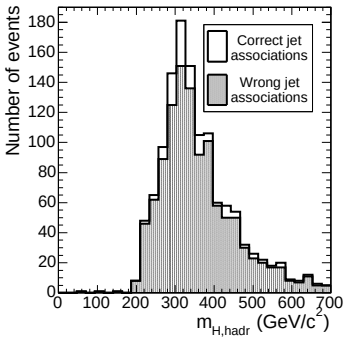


Figure 12: Reconstructed $\mathrm{H}^\pm$ mass with hadronically decaying top quark for the chosen jet association ($m_{\mathrm{H}^\pm} = 311\,\mathrm{GeV}$).

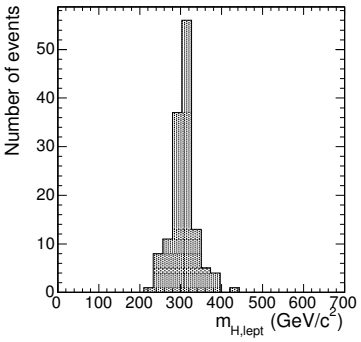


Figure 13: Reconstructed $\mathrm{H}^\pm$ mass with leptonically decaying top quark for correct jet associations ($m_{\mathrm{H}^\pm} = 311\,\mathrm{GeV}$).

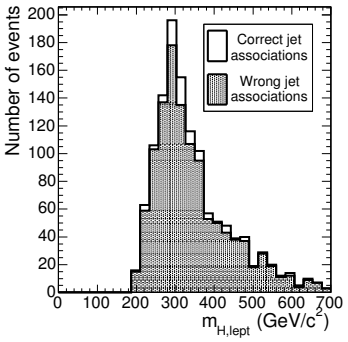


Figure 14: Reconstructed $\mathrm{H}^\pm$ mass with leptonically decaying top quark for the chosen jet association ($m_{\mathrm{H}^\pm} = 311\,\mathrm{GeV}$).

Due to the many possible jet associations and the ambiguity in the choice of the top quark candidate originating from the charged Higgs boson decay, the combinatorial background is very large. In Figures 12 and 14 the events for which jets are correctly associated to the decay products of the Higgs boson are shown on top of the combi-

natorial background. Because of this large combinatorial background no mass reconstruction is attempted, and the analysis is considered as a counting experiment. This way all events are retained and bad jet associations also contribute to the visibility of the signal.

4.3 Background suppression

To suppress the large $\mathrm{t\bar{t}b}/\mathrm{t\bar{t}j}$ background, observables were identified that have different properties for signal and background events. The following quantities were considered for the chosen jet association:

- observables associated with the $\mathrm{t\bar{t}}$ system: the transverse momentum $p_T(\mathrm{q}')$ of the softest jet q' from the $\mathrm{W}^{\pm}$ boson decay (Figure 15), and the χ^2 probability of the kinematic fit (Figure 16);
- the b-jet discriminator ζ_{b} for the b originating directly from the $\mathrm{H}^{\pm}$ decay (Figure 17);
- the ratio of the E_T of the sixth over the fifth jet, provided a sixth jet is found (Figure 18).

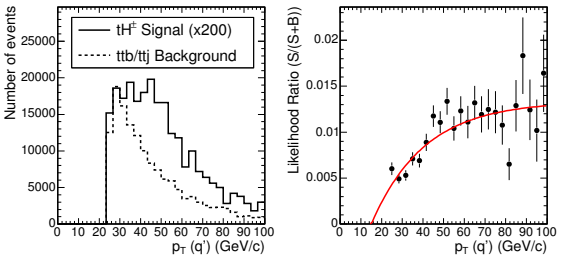


Figure 15: $p_T(\mathrm{q}')$ for $m_{\mathrm{H}^{\pm}} = 311\,\mathrm{GeV}$.

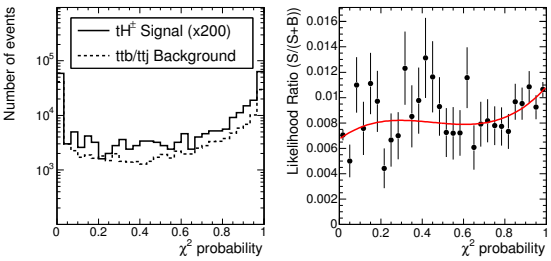


Figure 16: χ^2 probability for $m_{\mathrm{H}^{\pm}} = 311\,\mathrm{GeV}$.

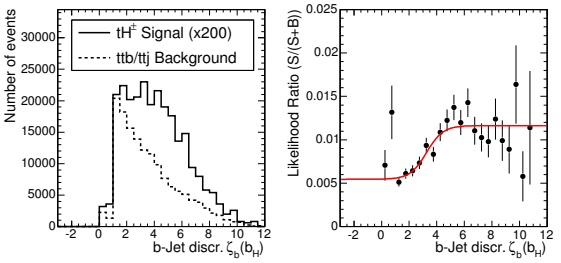


Figure 17: b-jet discriminator ζ_{b} of $\mathrm{b}_{\mathrm{H}^{\pm}}$ for $m_{\mathrm{H}^{\pm}} = 311\,\mathrm{GeV}$.

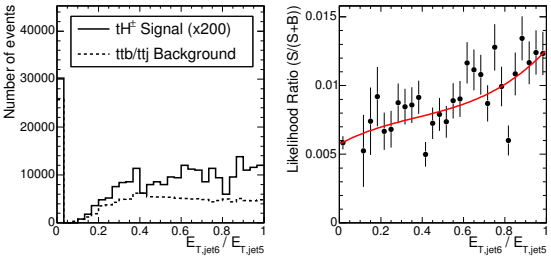


Figure 18: $E_{T,6}/E_{T,5}$ for $m_{\mathrm{H}^{\pm}} = 311\,\mathrm{GeV}$.

Using these observables, likelihood-ratio functions were constructed in a similar way as for the jet association and combined in a single ratio

$$\mathcal{L}_{\mathrm{SB}} = \prod_i \mathcal{L}_i, \tag{8}$$

shown in Figure 19 for both the signal and the background.

4.4 Statistical significance and discovery potential

The statistical significance of the observation of the signal is calculated using the S_{cP} program [24]. It translates the probability that a number of background events with Poissonian statistics mimicks an additional number of signal events into the number of Gaussian standard deviations that corresponds to this probability.

A cut on the combined likelihood ratio $\mathcal{L}_{\mathrm{SB}}$ increases the signal-to-background ratio. The significance of the signal observation is calculated as a function of this cut for an integrated luminosity of $30\,\mathrm{fb}^{-1}$, and its maximum, corresponding to an optimized cut on $\mathcal{L}_{\mathrm{SB}}$, is retained. In Figure 20 the statistical significance as a function of the cut on the combined likelihood ratio is shown.

With the knowledge of the maximal statistical significance for a certain integrated luminosity, the signal cross section for a 5σ discovery is derived for a given Higgs-boson mass. This corresponds to a minimal value of $\tan\beta$. Performing the described maximalization of the significance at different values of m_{A}, an optimized discovery

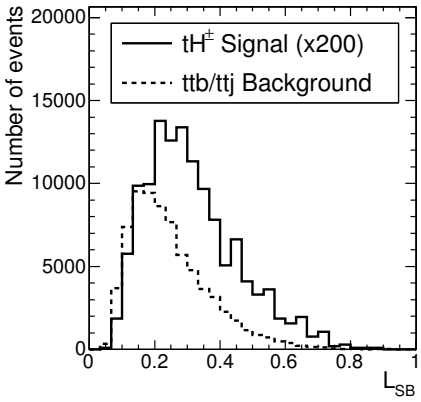


Figure 19: Distribution of the combined likelihood ratio $\mathcal{L}_{\mathrm{SB}}$ used to distinguish between signal and background.

contour is obtained in the MSSM $(m_{\mathrm{A}}, \tan\beta)$ plane, as shown in Figure 21 for an integrated luminosity of $30\,\mathrm{fb}^{-1}$ and $60\,\mathrm{fb}^{-1}$. On the same figure the exclusion limit at 95% confidence level for $30\,\mathrm{fb}^{-1}$ is drawn. Systematic uncertainties are not included.

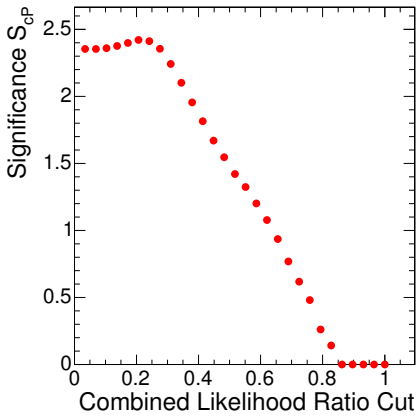


Figure 20: Statistical significance as a function of the cut on $\mathcal{L}_{\mathrm{SB}}$ for $m_{\mathrm{H}^{\pm}} = 311\,\mathrm{GeV}$ and $30\,\mathrm{fb}^{-1}$.

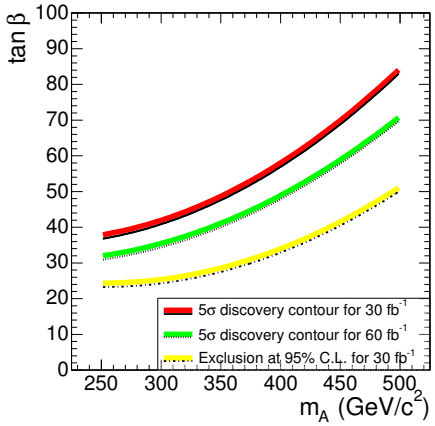


Figure 21: Discovery contours for the charged Higgs boson in the $\mathrm{H}^{\pm} \to \mathrm{tb}$ decay, applying 3 b tags, for $30\,\mathrm{fb}^{-1}$ and $60\,\mathrm{fb}^{-1}$. The discovery region lies above the curves. Also shown is the exclusion limit at 95% C.L. for $30\,\mathrm{fb}^{-1}$. Systematic uncertainties are not included.

4.5 Systematic uncertainties

The significance calculated above corresponds to the ideal case of perfect knowledge of the background cross section. In order to measure the background cross section and uncertainty from the data, a signal-free sample should be obtained. For this analysis, however, the signal and background are kinematically very similar. A possible way to estimate the background level from data is to impose the same selection cuts on the data as in Section 3, but to require one b-tagged jet less. Even with two b-tagged jets, the main background remains $\mathrm{t\bar{t}}$ production. After such a selection with only two b-tagged jets, the signal-to-background ratio is further reduced. It is then possible to calculate the expected number of background events plus its uncertainty, when tagging a third b jet.

For this method, a b-tagging efficiency and purity need to be measured from data. Making use of enriched b-jet samples from $\mathrm{t\bar{t}}$ events, the b-tagging efficiency in CMS is expected to be known to about 5% relative uncertainty with $10\,\mathrm{fb}^{-1}$ of integrated luminosity [25]. For mistag probabilities, such expectation estimations have not been performed yet. In the CDF [26] and DØ [27] experiments an uncertainty on the mistag probability of 10% was found for a secondary-vertex technique. Supposing an optimistic case where CMS obtains a 5% relative uncertainty on the mistag probability, then the uncertainty on the expected background level will at best be of the order of 5%.

Large theoretical uncertainties could also come into play using this method, like the ratio of $\mathrm{t\bar{t}} + \mathrm{b}$ events with real extra b jets and $\mathrm{t\bar{t}} + \mathrm{j}$ events with only jets from light quarks and gluons accompanying the top quarks.

The systematic effects on the discovery potential were evaluated, adding the systematic uncertainty in quadrature to the statistical contribution. The value of S/B has to be sufficiently large for the systematic contribution to the background uncertainty to be kept under control. Figure 22 shows the result of the significance calculation as a function of $\mathcal{L}_{\mathrm{SB}}$ for a systematic uncertainty on the background of only 1%. Comparing this with the case of perfect background knowledge in Figure 20, a large drop in significance for a low cut on $\mathcal{L}_{\mathrm{SB}}$ is observed, as expected from the large number of background events in this region.

Depending on the expected systematic uncertainty on the background level the maximal significance and its corresponding optimized cut on $\mathcal{L}_{\mathrm{SB}}$ was determined. In Figure 23 the discovery contours are plotted, for perfect knowledge of the $\mathrm{t\bar{t}b}/\mathrm{t\bar{t}j}$ cross section, for a 1% uncertainty and for a 3% uncertainty. From the above estimate of the systematic uncertainty on the number of background events, the conclusion is drawn that with this analysis, no sensitivity in this channel with triple b tagging is obtained in the MSSM parameter space during the low luminosity phase of LHC.

Assuming the same signal kinematics as for the MSSM charged Higgs-boson production, the required $\sigma(\mathrm{pp} \to \mathrm{tH^{\pm}X}) \times \mathrm{BR}(\mathrm{H^{\pm}} \to \mathrm{tb})$ for a five-sigma discovery by tagging three b jets is extracted. It is shown in Figure 24 as a function of m_{A}.

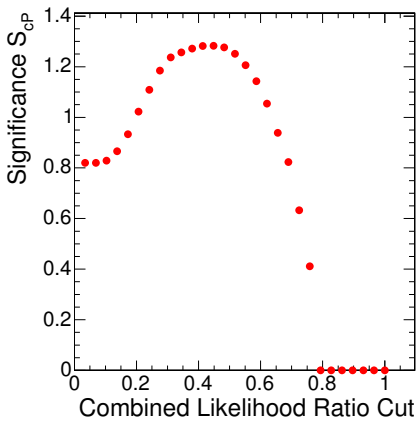


Figure 22: Significance as a function of the cut on $\mathcal{L}_{\mathrm{SB}}$ for $m_{\mathrm{H^{\pm}}} = 311\,\mathrm{GeV}$ and $30\,\mathrm{fb}^{-1}$, taking a systematic uncertainty of 1% on the background cross section into account.

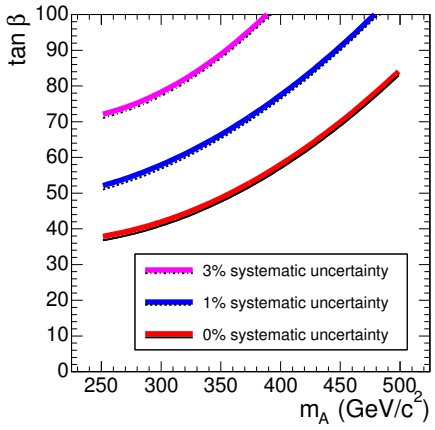


Figure 23: Discovery contour for the charged Higgs boson in the $\mathrm{H^{\pm}} \to \mathrm{tb}$ decay, applying 3 b tags, for $30\,\mathrm{fb}^{-1}$; systematic uncertainties on the background of 0%, 1% and 3% are taken into account. The discovery region lies above the curves.

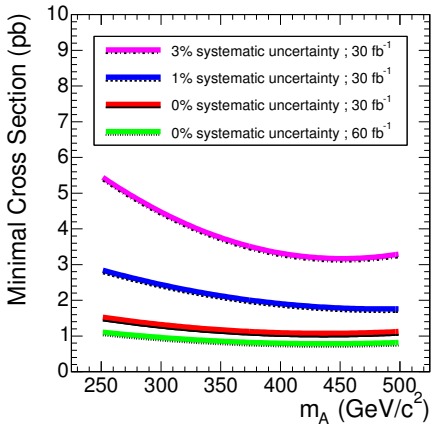


Figure 24: Required $\sigma(\mathrm{pp} \to \mathrm{tH^{\pm}X}) \times \mathrm{BR}(\mathrm{H^{\pm}} \to \mathrm{tb})$ for discovery of a signal with the same kinematics as in the MSSM, tagging 3 b jets, as a function of m_{A}.

5 Charged Higgs-Boson Search in the Four b-Jet Channel

5.1 Selection of the optimal jet association

When resolving the spectator b quark in the detector for the final state (4), six jets are expected in total, of which four b jets. These correspond to 360 possible jet associations. The b-tagging information reduces this number to 24. Several observables were identified to be able to choose the best jet association:

- observables associated with the output of the kinematic fit: the χ^2 probability of the kinematic fit (Figure 25), and the mass difference $m(\mathrm{t}_{\mathrm{hadr,fit}}) - m(\mathrm{t}_{\mathrm{hadr,rec}})$ between the hadronically-decaying top after and before the fit (Figure 26);
- the transverse momentum and pseudorapidity of the b jet from the $\mathrm{H}^{\pm}$ decay: $p_T(\mathrm{b}_{\mathrm{H}^{\pm}})$ (Figure 27) and $|\eta(\mathrm{b}_{\mathrm{H}^{\pm}})|$ (Figure 28);
- the b-jet discriminator ζ_{b} for the b jet originating directly from the $\mathrm{H}^{\pm}$ decay (Figure 29);
- the b-jet discriminator ζ_{b} for the jet associated with the spectator b quark (Figure 30);
- the combined b-identification probability

$$P_{2\mathrm{b}} = \frac{P_{\mathrm{b}}(\mathrm{b}_1) + P_{\mathrm{b}}(\mathrm{b}_2)}{P_{\mathrm{b}}(\mathrm{b}_1) + P_{\mathrm{b}}(\mathrm{b}_2) + P_{\mathrm{b}}(\mathrm{j}_1) + P_{\mathrm{b}}(\mathrm{j}_2)},$$

  obtained combining the b-jet probabilities P_{b} for all the jets coming from top decays (Figure 31).

The jet associations used to obtain the distributions for these observables include the ones disfavoured by the b tagging, except for the associations for which the kinematic fit fails. Because of the small number of simulated events the distributions were made using events from the same samples with only two b tags applied. A likelihood ratio function $\mathcal{L}_i$ was obtained for each of the observables by calculating bin-by-bin the ratio $S/(S+B)$, where S and B denote respectively the number of correct and wrong jet associations, and fitting the resulting distribution. For each jet association the combined likelihood ratio $\mathcal{L}_{\mathrm{sol}}$ value is calculated with (7). The jet pairing with the highest value for $\mathcal{L}_{\mathrm{sol}}$ is chosen as the best possible jet association.

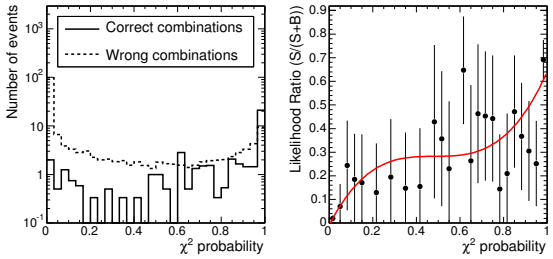


Figure 25: χ^2 probability for $m_{\mathrm{H}^{\pm}} = 311\,\mathrm{GeV}$.

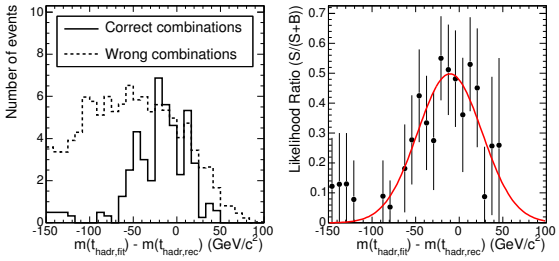


Figure 26: $m(\mathrm{t}_{\mathrm{hadr,fit}}) - m(\mathrm{t}_{\mathrm{hadr,rec}})$ for $m_{\mathrm{H}^{\pm}} = 311\,\mathrm{GeV}$.

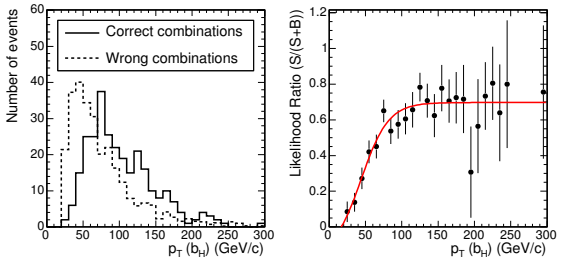


Figure 27: $p_T(\mathrm{b}_{\mathrm{H}^{\pm}})$ for $m_{\mathrm{H}^{\pm}} = 311\,\mathrm{GeV}$.

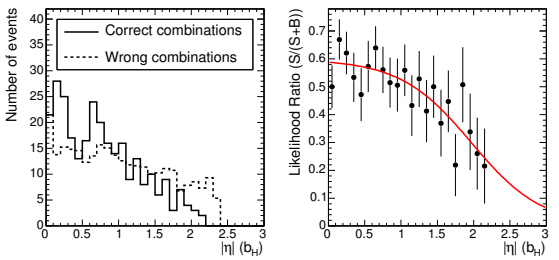


Figure 28: $|\eta(\mathrm{b}_{\mathrm{H}^{\pm}})|$ for $m_{\mathrm{H}^{\pm}} = 311\,\mathrm{GeV}$.

5.2 Background suppression

In the case of the final state with four b quarks, the background consists of the irreducible, but relatively small, $\mathrm{t\bar{t}b\bar{b}}$ process, and the reducible, but very large, $\mathrm{t\bar{t}jj}$ process. Therefore the background suppression mostly relies

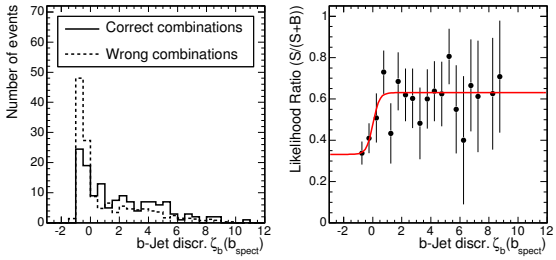


Figure 29: b-jet discriminator ζ_b of b_{spect} for $m_{H^\pm} = 311\,\text{GeV}$.

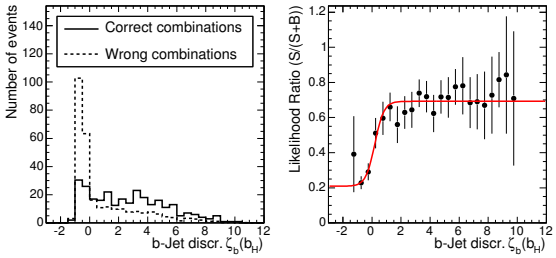


Figure 30: b-jet discriminator ζ_b of $b_{H^\pm}$ for $m_{H^\pm} = 311\,\text{GeV}$.

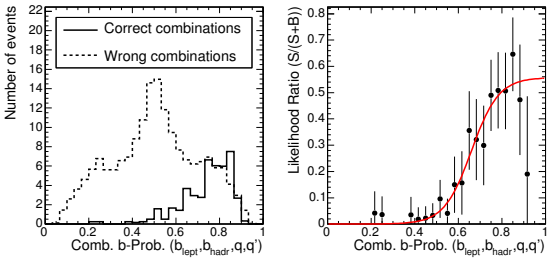


Figure 31: Combined b probability P_{2b} for $m_{H^\pm} = 311\,\text{GeV}$.

on the b tagging of the jets accompanying the $t\bar{t}$ pair. Other observables, like the momentum and pseudorapidity of the top candidates and of the extra jets, present some additional differences between signal and background, especially for large Higgs boson masses. Owing to the low statistics, however, they could not be shown to have a significant impact on background rejection, and are hence not used in the following. In light of the limited statistics, the most effective variable to discriminate signal and background was found to be the sum of the b-tag discriminators $\zeta_b(b_{H^\pm})$ and $\zeta_b(b_{spect})$ of the b jets associated to the Higgs boson decay and to the spectator jet respectively. The likelihood variable $\mathcal{L}_{SB}$, defined as $(2 + \zeta_b(b_{H^\pm}) + \zeta_b(b_{spect}))/26$ such that $0 < \mathcal{L}_{SB} < 1$, is shown in Figure 32.

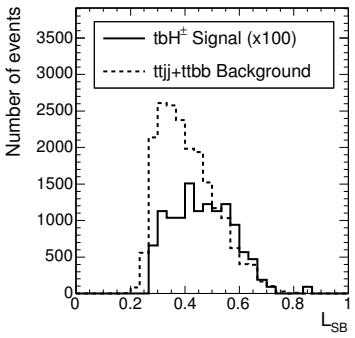


Figure 32: Distribution of the variable $\mathcal{L}_{SB}$ used to distinguish between signal and background.

5.3 Statistical significance and discovery potential

By cutting on $\mathcal{L}_{SB}$ it is possible to increase the signal-to-background ratio. Calculating the significance of the signal as a function of this cut, the point of maximal significance is determined for a given integrated luminosity, taken to be $30\,\text{fb}^{-1}$. To accommodate the small signal statistics and large scale factor (~ 8) needed for the background simulation, the number of signal events and the corresponding number of background events after a cut on $\mathcal{L}_{SB}$ was parametrized, as shown in Figure 33. This parametrized dependency was used when calculating the significance as a function of $\mathcal{L}_{SB}$ rather than the direct dependence on $\mathcal{L}_{SB}$ of the number of signal and background events. In Figure 34 the resulting statistical significance as a function of the cut on $\mathcal{L}_{SB}$ is shown.

With the knowledge of the maximal statistical significance for a certain integrated luminosity, the signal cross section for a 5σ discovery is derived for a given Higgs-boson mass. This corresponds to a minimal value of $\tan\beta$.

Performing the described maximalization of the significance at different values of m_{A}, an optimized discovery contour is obtained in the MSSM $(m_{\mathrm{A}}, \tan\beta)$ plane, as shown in Figure 35 for an integrated luminosity of $30\,\mathrm{fb}^{-1}$ and $60\,\mathrm{fb}^{-1}$. On the same figure the exclusion limit at 95% confidence level for $30\,\mathrm{fb}^{-1}$ is drawn. Systematic uncertainties are not included.

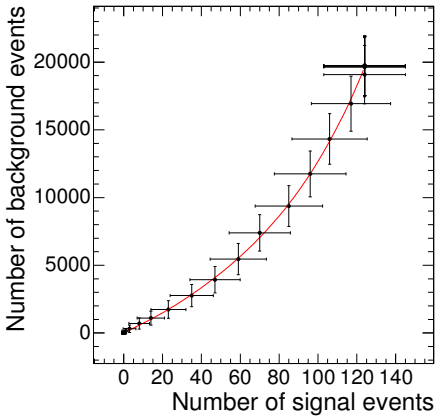


Figure 33: Number of background events as a function of the number of signal events, for various cuts on $\mathcal{L}_{\mathrm{SB}}$.

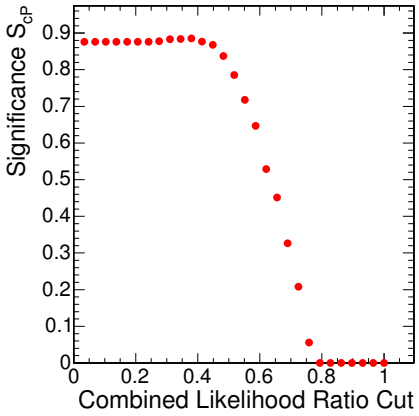


Figure 34: Statistical significance as a function of the cut on $\mathcal{L}_{\mathrm{SB}}$ for $m_{\mathrm{H}^{\pm}} = 311\,\mathrm{GeV}$ and $30\,\mathrm{fb}^{-1}$.

5.4 Systematic uncertainties

The significance calculated above corresponds to the ideal case of perfect knowledge of the background cross section. In order to measure the background cross section level and uncertainty a similar technique can be used as in the case of triple b tagging. An optimistic uncertainty of 5% on b-jet identification and mistag probability measurements is assumed, neglecting possible theoretical uncertainties.

As in the three b-jet case, the significance calculation was repeated for a hypothetical systematic uncertainty on the level of the background of 1% and 3%. The resulting significance as a function of $\mathcal{L}_{\mathrm{SB}}$ for 1% uncertainty is shown in Figure 36. The corresponding discovery contours are plotted in Figure 37. A value of 3% is no longer visible in the figure. From the above estimate of the systematic uncertainty on the number of background events, the conclusion is drawn that with the presented analysis, no sensitivity in this channel with four b-tagged jets is obtained in the MSSM parameter space during the low luminosity phase of LHC.

Assuming the same signal kinematics as for the MSSM charged Higgs-boson production, the required $\sigma(\mathrm{pp} \to \mathrm{tH}^{\pm}\mathrm{X}) \times \mathrm{BR}(\mathrm{H}^{\pm} \to \mathrm{tb})$ for a five-sigma discovery by tagging 4 b jets is extracted. It is shown in Figure 38 as a function of m_{A}.

6 Conclusion

The observation prospects for a charged Higgs boson produced in the $\mathrm{pp} \to \mathrm{tH}^{\pm}\mathrm{X}$ channel are discussed for a search relying on three or four b-tagged jets. Cascades from heavy sparticles and decays into lighter SUSY particles were not considered. The only important background for this channel, $\mathrm{t\bar{t}}$ production with additional real or mistagged b jets, is large and kinematically similar to the signal.

A detailed event reconstruction and selection is performed with parametrized or full CMS detector simulation in the three or four b-jet search, respectively. Likelihood-ratio methods are applied to choose the best jet association and to suppress the background. The statistical significance is estimated and 5σ-discovery contours in the $(m_{\mathrm{A}}, \tan\beta)$ plane are constructed. The influence of a systematic uncertainty on the background knowledge is evaluated, and is found to have a large effect on the visibility of the signal in the MSSM, due to the very small signal-to-background ratio, in both the three and four b-tag case.

In conclusion, no sensitivity is obtained in the MSSM parameter space with this analysis, due to the large background and resulting large effects of systematic uncertainties on the background knowledge. Assuming the same signal kinematics as for the MSSM case, the required $\sigma(\mathrm{pp} \to \mathrm{tH}^{\pm}\mathrm{X}) \times \mathrm{BR}(\mathrm{H}^{\pm} \to \mathrm{tb})$ for a five-sigma discovery is extracted for the final states with three or four b-tagged jets.

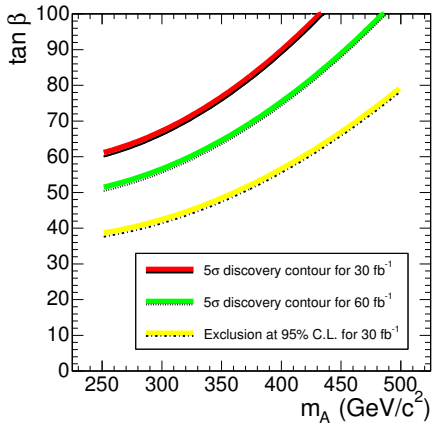


Figure 35: Discovery contour for the charged Higgs boson in the $\mathrm{H}^{\pm} \to \mathrm{tb}$ decay, applying 4 b tags, for $30\,\mathrm{fb}^{-1}$ and $60\,\mathrm{fb}^{-1}$. The discovery region lies above the curves. Also shown is the exclusion limit at 95% C.L. for $30\,\mathrm{fb}^{-1}$. Systematics are not included.

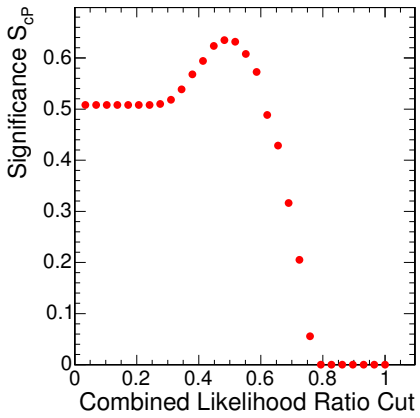


Figure 36: Significance as a function of the cut on $\mathcal{L}_{\mathrm{SB}}$ for $m_{\mathrm{H}^{\pm}} = 311\,\mathrm{GeV}$ and $30\,\mathrm{fb}^{-1}$, taking a systematic uncertainty of 1% on the background cross section into account.

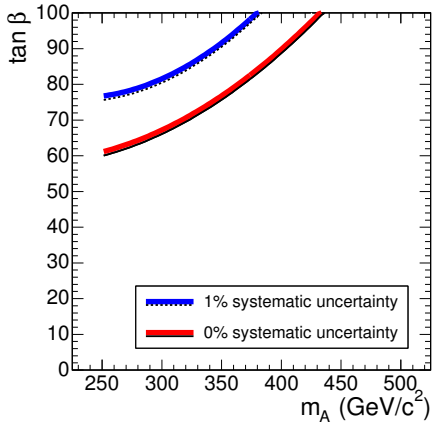


Figure 37: Discovery contour for the charged Higgs boson in the $\mathrm{H}^{\pm} \to \mathrm{tb}$ decay, applying 4 b tags, for $30\,\mathrm{fb}^{-1}$; systematic uncertainties on the background of 0%, 1% and 3% are taken into account. The discovery region lies above the curves.

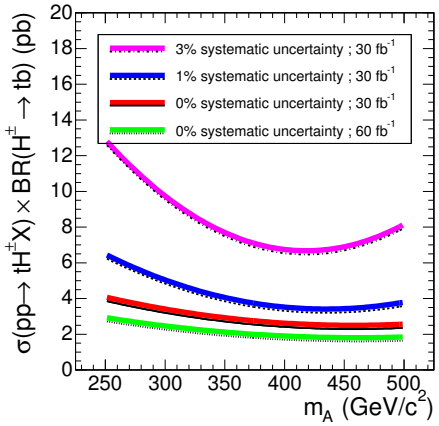


Figure 38: Required $\sigma(\mathrm{pp} \to \mathrm{tH}^{\pm}\mathrm{X}) \times \mathrm{BR}(\mathrm{H}^{\pm} \to \mathrm{tb})$ for discovery of a signal with the same underlying physics, tagging 4 b jets, as a function of m_{A}.

Acknowledgments

We would like to thank Sasha Nikitenko, Fabio Maltoni, Tilman Plehn, Johan Alwall, Christian Weiser, Alexander Schmidt and Susanna Cucciarelli for the many helpful discussions, and Salvatore Mele and Martijn Mulders for the detailed reading of the manuscript.

References

[1] J. F. Gunion, H. E. Haber, G. L. Kane, and S. Dawson, *The Higgs Hunters' Guide*. Addison-Wesley, Reading, MA, 1990.

[2] S. P. Martin, *A Supersymmetry Primer*, hep-ph/9709356.

[3] V. D. Barger, R. J. N. Phillips, and D. P. Roy, Phys. Lett. **B324** (1994) 236–240.

[4] J. F. Gunion, Phys. Lett. **B322** (1994) 125–130.

[5] S. Moretti and D. P. Roy, Phys. Lett. **B470** (1999) 209–214.

[6] D. J. Miller, S. Moretti, D. P. Roy, and W. J. Stirling, Phys. Rev. **D61** (2000) 055011.

[7] D. Cavalli *et al.*, *The Higgs Working Group: Summary Report*, hep-ph/0203056.

[8] S. Lowette, P. Vanlaer, and J. Heyninck, *Heavy Charged MSSM Higgs Bosons in the $H^{\pm} \to tb$ Decay in CMS*, CMS NOTE 2004/017, 2004.

[9] P. Salmi, R. Kinnunen, and N. Stepanov, *Prospects of Detecting Massive Charged Higgs from Hadronic Decay $H^{\pm} \to tb$ in CMS*, CMS NOTE 2002/024, 2002.

[10] T. Sjostrand *et al.*, Comput. Phys. Commun. **135** (2001) 238–259.

[11] J. Alwall, *MATCHIG: A Program for Matching Charged Higgs Boson Production at Hadron Colliders*, hep-ph/0503124.

[12] T. Plehn, Phys. Rev. **D67** (2003) 014–018.

[13] A. Djouadi, J. Kalinowski, and M. Spira, Comput. Phys. Commun. **108** (1998) 56–74.

[14] J. Heyninck, J. D'Hondt, and S. Lowette, *Top Quark Mass Measurement in Single Leptonic $t\bar{t}$ Events*, CMS NOTE 2006/066, 2006.

[15] F. Maltoni and T. Stelzer, JHEP **02** (2003) 027.

[16] CompHEP Collaboration, E. Boos *et al.*, Nucl. Instrum. Meth. **A534** (2004) 250–259.

[17] M. L. Mangano, M. Moretti, F. Piccinini, R. Pittau, and A. D. Polosa, JHEP **07** (2003) 001.

[18] GEANT4 Collaboration, S. Agostinelli *et al.*, Nucl. Instrum. Meth. **A506** (2003) 250–303.

[19] J. D'Hondt, S. Lowette, and J. Heyninck, *Electron and Muon Reconstruction in Single Leptonic $t\bar{t}$ Events*, CMS NOTE 2006/024, 2006.

[20] A. Heister, O. Kodolova, V. Konoplianikov, S. Petrushanko, J. Rohlf, C. Tully, and A. Ulyanov, *Measurement of Jets with the CMS Detector at the LHC*, CMS NOTE 2006/036, 2006.

[21] C. Weiser, *A Combined Secondary Vertex Based B-Tagging Algorithm in CMS*, CMS NOTE 2006/014, 2006.

[22] CMS Collaboration, *Technical Design Report: The TriDAS Project, Volume II: Data Acquisition & High Level Trigger*, CERN/LHCC/2002-26, 2002.

[23] J. D'Hondt, S. Lowette, O. Buchmuller, S. Cucciarelli, F.-P. Schilling, M. Spiropulu, S. Paktinat Mehdiabadi, D. Benedetti, and L. Pape, *Fitting of Event Topologies with External Kinematic Constraints in CMS*, CMS NOTE 2006/023, 2006.

[24] *Programs for Significance Calculations in CMS*, http://cmsdoc.cern.ch/∼bityukov/.

[25] J. D'Hondt, S. Lowette, J. Heyninck, and P. Vanlaer, *Offline Calibration of b-Jet Identification Efficiencies*, CMS NOTE 2006/013, 2006.

[26] CDF Collaboration, T. Affolder *et al.*, Phys. Rev. **D64** (2001) 032002.

[27] D0 Collaboration, V. M. Abazov *et al.*, Phys. Lett. **B626** (2005) 35–44.